General–to–Specific Model Selection Procedures
for Structural Vector Autoregressions

Hans-Martin Krolzig [*]

Department of Economics and Nuffield College, Oxford University.
hans-martin.krolzig@nuffield.oxford.ac.uk

First version: December 2001

This version: March 2003

Abstract

Structural vector autoregressive (SVAR) models have emerged as a dominant research strategy in empirical macroeconomics, but suffer from the large number of parameters employed and the resulting estimation uncertainty associated with their impulse responses. In this paper we propose general-to-specific model selection procedures to overcome these limitations. After showing that single-equation procedures are efficient for the reduction of the SVAR, but generally not for the reduction of its reduced form, the proposed reduction procedure is computer-automated using *PcGets* and its small-sample properties are evaluated in a realistic Monte Carlo experiment. The model selection procedure is shown to recover the DGP specification from a large unrestricted SVAR model with controlled size and power. The impulse responses generated by the selected SVAR are compared to those of the unrestricted and reduced VAR and found to be more precise and accurate. The proposed reduction strategy is then applied to the US monetary system considered by Christiano, Eichenbaum and Evans (1996). Although the selection process is hampered by the misspecification of the unrestricted VAR, the results are consistent with the Monte Carlo and question the validity of the impulses responses generated by the full system.

JEL Classification: C51, C32, E52.

Keywords: Model selection; Impulse responses; Vector autoregression; Structural VAR; Causal order; Data mining.

1 Introduction

Over the last two decades, vector autoregressive (VAR) models have emerged as an important research tool for the empirical analysis of macroeconomic time series, partly because of the critique in Sims (1980) of traditional macro-econometric modelling. VARs have been widely exploited for the description of numerous macroeconomic data sets, offering fruitful insight on the interrelations between economic variables. The popularity of VARs is due to various advantages of the approach: First, the flexibility of the VAR framework in producing econometric models with useful descriptive characteristics, within which statistical tests of economically meaningful hypothesis can be executed. Secondly, the ease of the approach, as econometric models can be formulated and data characterized without having to invoke economic theory to restrict the dynamic relations between variables. Thirdly, the characterization of macroeconomic models, as many completely specified economic models give rise to VAR representations as the reduced form of the variables of the system. Fourthly, the compatibility with quite a wide variety of hypotheses regarding the formation of expectations. Despite these advantages,

---

1

it is widely acknowledged in the literature that, in general, the innovations of VARs are not identified with the underlying structural errors due to the correlation of residuals across equations as in the case of instantaneous causality,. Therefore, the impulse responses generated by such a VAR do not possess a structural interpretation. While there is no unique best way to deal with this problem, a popular way of overcoming the problem, since Sims (1980), is the transformation of the residuals to orthogonal form by triangulating the system, which involves a causal ordering of the variables. The transformed VAR allows the interpretation of the evolution of the system as a function of the orthogonalized innovations in the variables of the system. A related approach to respond to the problem of interpreting VARs has been the development of structural vector autoregressions (SVARs), which introduce 'theoretical' restrictions to identify the underlying shocks. The SVAR approach tends to impose just enough restrictions to permit a coherent interpretation of the shocks to the system. In this paper we focus on SVARs made recursive in contemporaneous variables. In contrast to SVARs in the spirit of Blanchard and Quah, 1989, identification is achieved by short-run restrictions specifying the causal ordering of the variables in the system.

SVAR models suffer from the fact that they rarely impose any restrictions upon the dynamics in their implied structural equations. For example, the relatively small structural VAR models of Bernanke and Blinder (1992) and Sims (1992) for the US economy have the distinctive feature that each structural equation is saturated with lagged variables *i.e.* the dynamics are essentially unrestricted. In such a just-identified SVAR, the number of parameters grows with the square of the number of variables and quickly exhausts the degree of freedom ("curse of dimensionality"). Due to the large number of model parameters, the structural equations of the SVAR are not only estimated imprecisely, but also hard to interpret. These considerations point to the need for reductions of the systems which involve the utilization of exclusion restrictions upon the dynamics contained in each structural equation, so as to allow for easier interpretation of the system. It seems sensible to employ model reduction procedures, which so far have been mainly used in the single-equation framework. For the construction of a recursive SVAR of the Australian economy, Dungey and Pagan (2000) employ a modelling approach combining simplifications based on statistical tests and economic considerations. In this paper, we present a reduction process proceeding by imposing zero restrictions according to the outcome of statistical tests and abstract from the use of economic theory for the derivation of overidentifying restrictions

The existing literature on VAR model selection has mainly focused on the selection of lag order, p, of an otherwise unrestricted reduced-form VAR. In these selection procedures, a model is usually selected by an information criterion which penalizes the likelihood function for the number of parameters. Lütkepohl (1991) discusses various strategies for selecting subset VAR models (*i.e.*, VARs with zero constraints on the coefficients), which are based on the optimization of a specified model selection criterion for a given maximal order of the VAR including full search, search over complete VAR matrices, top-down and bottom-up specification of the distributed lag lengths etc. Brüggemann and Lütkepohl (2000) consider step-wise-regression-type single-equation reduction paths where the critical value is chosen such that an acceptance of the null hypothesis guarantees a marginal increase in a given information criterion.

In this paper, we propose 'General-to-specific' (*Gets*) model selection procedures for SVAR models designed to overcome the limitations of just-identified SVAR models by reducing the number of required parameters. We will argue that the reduction and identification of SVAR models is a natural area for the application of *Gets* reduction procedures. The proposed reduction process is designed to ensure that the reduced SVAR model will convey all the information embodied in the unrestricted SVAR. This is achieved by a joint selection and diagnostic testing process: starting from the unrestricted, congruent general model, standard testing procedures are used to eliminate statistically-insignificant variables, with

diagnostic tests checking the validity of reductions, ensuring a congruent final selection. By reducing the complexity of the just-identified SVAR and simultaneously ensuring that the reduced SVAR will convey all the information embodied in the unrestricted VAR, the selected simpler, more compact model provides an improved statistical description of the economic world (see Hendry, 2000, for an overview of the so-called 'LSE' methodology). A further merit of this modelling approach is that it systematically checks for the presence of statistical misspecifications. In a substantial number of papers, the restrictions used to identify the VAR are imposed without establishing the congruence with the data (*e.g.* absence of regime changes and the constancy of the estimated coefficients: see, *inter alia*, Hendry and Mizon, 2000). For the economic interpretation of the SVAR and the generality of the derived impulse responses, it is therefore essential to ensure the congruence of the assumptions made.

The recent developments in automatic model selection initiated by Hoover and Perez (1999) suggests that the operational characteristics of some computer-automated model selection algorithms are excellent across a wide range of states of nature. For the computer implementation of the proposed model selection procedure for SVAR, we naturally focus on *PcGets* developed by Hendry and Krolzig (2001). *PcGets* automates general-to-specific (*Gets*) modelling for linear, dynamic, single-equation models based on the outlined theory of reduction. For a more detailed description of the algorithm see Hendry and Krolzig (2003). In Krolzig (2001) and Brüggemann, Krolzig and Lütkepohl (2002), computer-automated model selection algorithms such as *PcGets* were examined for the reduction of reduced-form VAR models and found to deliver reasonable results. It will be shown that single-equation procedures such as *PcGets* are efficient for the reduction of reduced-form VAR models under the condition of conditional independence, and for recursive structural VAR models in general.

The structure of the paper is as follows: The following section (§2) defines the reduced-form and structural VAR model studied in this paper. §3 discusses the theoretical properties of *Gets* reduction procedures for the reduction of VAR processes. The small sample characteristics, foremost its selection properties and the precision and accuracy of the generated impulse-responses, of the proposed *Gets* model-selection procedure as implemented by *PcGets* are then investigated by simulation. In the realistic Monte Carlo experiment of §4, the data generating process is an over-identified trivariate SVAR(1) and the general unrestricted model is an SVAR(5) or VAR(5). An empirical illustration with a US monetary system based on Christiano *et al.* (1996) evaluating the usefulness of the proposed approach for the analysis of large macroeconomic data sets, follows in §5. In §6 we outline generalizations of the proposed modelling approach (i) for the selection of the causal order of the variables of the system, (ii) for the reduction of cointegrated time series models and (iii) for the search of simplifications as well as omitted variables in identified simultaneous equation models. Finally §7 concludes.

2 The vector autoregressive model

2.1 The reduced-form VAR

The basic model considered in the following is a vector autoregression possibly including deterministic terms and with independent Gaussian errors: the n-dimensional time series vector $\mathbf{y}_t$ is generated by a stationary vector autoregressive process of order p, denoted VAR(p) model,

$$\mathbf{y}_t = \boldsymbol{\nu} + \sum_{i=1}^{p} \mathbf{A}_i \mathbf{y}_{t-i} + \boldsymbol{\varepsilon}_t, \tag{1}$$

where $t = 1, \ldots, T$, the $\mathbf{A}_i$ and $\boldsymbol{\nu}$ are coefficient matrices and the initial values of $\mathbf{Y}_0 = (\mathbf{y}_0, \ldots, \mathbf{y}_{1-p})$ are fixed. Also, $\det\left(\mathbf{I}_K - \mathbf{A}_1 z - \ldots - \mathbf{A}_p z^p\right) \neq 0$ for $|z| \leq 1$. The innovation process $\boldsymbol{\varepsilon}_t$ is an

unobservable Gaussian zero-mean vector white noise process with a time-invariant positive-definite variance-covariance matrix $\mathsf{E}[\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'] = \boldsymbol{\Sigma}$ is given by:

$$\boldsymbol{\varepsilon}_t \sim \mathsf{NID}(\mathbf{0}, \boldsymbol{\Sigma}). \tag{2}$$

The infinite-order vector moving-average representation of the VAR in (1) is

$$\mathbf{y}_t = \boldsymbol{\mu} + \sum_{j=0}^{\infty} \boldsymbol{\Psi}_j \boldsymbol{\varepsilon}_{t-j}, \tag{3}$$

where $\boldsymbol{\mu} = \left(\sum_{i=1}^{p} \mathbf{A}_i\right)^{-1} \boldsymbol{\nu}$ and $\boldsymbol{\Psi}(L) = \left(\mathbf{I} - \sum_{j=i}^{p} \mathbf{A}_i L^i\right)^{-1}$, such that for VAR(1) processes, $\boldsymbol{\Psi}_j = \mathbf{A}^j$. The (k,l)-th element $\psi_{kl,j}$ of the MA matrix $\boldsymbol{\Psi}_j$ can be interpreted as the reaction of variable k in response to a unit shock in variable l, j periods ago.

The assumption that the shocks occur only in one of the variables, as implicitly made in this type of impulse response analysis, is fully justified under conditional independence, but problematic if the residuals are correlated. In the later case, the VAR can be readily transformed to interpret the evolution of the system as a function of orthogonalized innovations in any of the variables. Define $\boldsymbol{\eta}_t^* = \mathbf{P}^{-1}\boldsymbol{\varepsilon}_t$ by decomposing $\boldsymbol{\Sigma}$ as $\boldsymbol{\Sigma} = PP'$, where $\mathbf{P}$ is a lower triangular matrix, such that $\boldsymbol{\eta}_t^* \sim \mathsf{NID}(\mathbf{0}, \mathbf{I}_K)$. The orthogonalized vector moving average representation is given by:

$$\mathbf{y}_t = \boldsymbol{\mu} + \sum_{j=0}^{\infty} \boldsymbol{\Psi}_j \mathbf{P}\mathbf{P}^{-1}\boldsymbol{\varepsilon}_{t-j} = \boldsymbol{\mu} + \sum_{j=0}^{\infty} \boldsymbol{\Phi}_j^* \boldsymbol{\eta}_{t-j}^*, \tag{4}$$

where $\boldsymbol{\Phi}_0^* = \mathbf{P}$ and $\boldsymbol{\Phi}_j^* = \boldsymbol{\Psi}_j \mathbf{P}$. It is a well-known fact that orthogonalized impulse-responses, which are based on a Choleski decomposition of the variance-covariance matrix of the reduced-form VAR, are not invariant against changes in the (causal) ordering of the variables.

2.2 The structural VAR

The type of structural vector autoregressive (SVAR) processes considered in this paper is:

$$\mathbf{B}\mathbf{y}_t = \boldsymbol{\delta} + \sum_{i=1}^{p} \boldsymbol{\Gamma}_i \mathbf{y}_{t-i} + \boldsymbol{\eta}_t, \tag{5}$$

where $\mathbf{B}$, $\boldsymbol{\Gamma}_i$ and $\boldsymbol{\delta}$ are coefficient matrices and the innovation process $\boldsymbol{\eta}_t$ is an unobservable Gaussian zero-mean vector white noise process with a time-invariant diagonal variance-covariance matrix $\mathsf{E}[\boldsymbol{\eta}_t\boldsymbol{\eta}_t'] = \boldsymbol{\Omega}$:

$$\boldsymbol{\eta}_t \sim \mathsf{NID}(\mathbf{0}, \boldsymbol{\Omega}). \tag{6}$$

The SVAR in equation 5 can be considered a particular simultaneous equation model in the spirit of the *Cowles approach*. Particularly, it is a recursive system of the sort proposed by Wold (1949) and Strotz and Wold (1960) and closely related to the concept of causal ordering introduced by Simon (1953).

To recover the structural parameters it is useful to consider the structural VAR in equation (5) in its reduced-form. The relation to the VAR in (1) is given by:

$$\begin{aligned} \boldsymbol{\Sigma} &= \mathbf{B}^{-1}\boldsymbol{\Omega}\mathbf{B}^{-1\prime}, \\ \mathbf{A}_i &= \mathbf{B}^{-1}\boldsymbol{\Gamma}_i \qquad \text{for } i = 1, \ldots, p, \\ \boldsymbol{\nu} &= \mathbf{B}^{-1}\boldsymbol{\delta}. \end{aligned} \tag{7}$$

The uniqueness of the model for the given structure, which guarantees the estimableness of the structural parameters, is ensured by the following set just-identifying restrictions:

$$
\begin{array}{rcll}
\omega_{ij} & = & 0 & \text{for } i \neq j, \\
\beta_{ii} & = & 1 & \text{for } i = 1, \ldots, n, \\
\beta_{ij} & = & 0 & \text{for } i < j.
\end{array}
\tag{8}
$$

The infinite-order structural vector moving-average representation results from (3) by $\varepsilon_{t-j} = \mathbf{B}^{-1}\boldsymbol{\eta}_{t-j}$ as:

$$
\mathbf{y}_t = \boldsymbol{\mu} + \sum_{j=0}^{\infty} \boldsymbol{\Psi}_j \mathbf{B}^{-1} \boldsymbol{\eta}_{t-j} = \boldsymbol{\mu} + \sum_{j=0}^{\infty} \Phi_j \boldsymbol{\eta}_{t-j},
\tag{9}
$$

where $\boldsymbol{\Phi}(L) = \left(\mathbf{B} - \sum_{i=1}^{p} \boldsymbol{\Gamma}_i L^i\right)^{-1}$ with $\boldsymbol{\Phi_0} = \mathbf{B}^{-1}$. For a given causal ordering of the variables in the SVAR, the representation (9) differs from (4) only by the missing adjustment for the standard errors of the η_{kt}. In other words, the relation to (4) is given by $\boldsymbol{\eta}_t^* = \boldsymbol{\Omega}^{-\frac{1}{2}}\boldsymbol{\eta}_t$.

3 General-to-specific reductions procedures for VAR models

VAR modelling is a natural area for the application of *Gets*: The unrestricted VAR(p) model constitutes the general unrestricted model (GUM) defining the model space to be searched for the unknown data generating process (DGP), which as we presuppose is a subset of the unrestricted VAR. Such systems can be analyzed one equation at a time, since every equation has the same set of regressors, but each variable is the regressand in turn.

The considered model selection procedure is of the form

$$
\xi : \boldsymbol{\chi} \to \boldsymbol{\Xi} : \mathbf{Y} \to \tilde{\boldsymbol{\xi}} = \xi(\mathbf{Y}),
\tag{10}
$$

where $\boldsymbol{\chi}$ is the observation space, $\boldsymbol{\Xi}$ is the model space, which collects all subsets models of unrestricted VAR, and $\mathbf{Y}$ is the observed sample. The selection problem consists of inclusion versus exclusion decisions for each coefficient of the full VAR and results in the binary selection vector $\tilde{\boldsymbol{\xi}} \in \boldsymbol{\Xi} = \{1, 0\}^n$ with ones signaling inclusion and zeros elimination of the coefficient, n is the number of coefficients in the model. For a reduced-form VAR without deterministic terms we have that $n = K^2 p$. Thus the dimension of model space, dim($\boldsymbol{\Xi}$) is $2^{K^2 p}$. We assume that the unrestricted VAR is congruent: the model space is consistent and includes the true selection vector $\boldsymbol{\xi}^{\mathrm{DGP}} \in \boldsymbol{\Xi}$.

The *Gets* reduction process relies on a classical, sequential-testing approach (see, *inter alia*, Hendry, 1995 and 2000) . Different critical values are set for multiple and single selection tests, and for diagnostic tests. Denote by $\boldsymbol{\eta}$ the vector of significance levels for the misspecification tests (diagnostics) and by $\boldsymbol{\alpha}$ the vector of significance level for the various selection tests. During the specification search, the current specification is simplified only if no diagnostic test rejects its null. This corresponds to a likelihood-based model evaluation, where the likelihood function accepts the probability density function of model $\boldsymbol{\xi}$, only if the sample information coheres with the underlying assumptions of the model itself, *i.e.* $\min\left(\tilde{\boldsymbol{\eta}}(\mathbf{Y}; \tilde{\boldsymbol{\theta}}_{\boldsymbol{\xi}}) - \boldsymbol{\eta}\right) < 0$, where the vector of diagnostic test statistics p-values, $\tilde{\boldsymbol{\eta}}(\mathbf{Y}; \tilde{\boldsymbol{\theta}}_{\boldsymbol{\xi}})$, is evaluated at the maximum likelihood estimate $\widetilde{\boldsymbol{\theta}}_{\boldsymbol{\xi}}$ under model $\boldsymbol{\xi}$, and mapped into its marginal rejection probabilities.

Since jointly selecting and diagnostic testing eludes theoretical analysis, we approximate the *Gets* reduction process by:

$$
\tilde{\boldsymbol{\xi}} = \arg\max_{\boldsymbol{\xi} \in \boldsymbol{\Xi}^{\mathbf{c}}} \left\{ \max_{\boldsymbol{\theta}} \left[\frac{1}{T} \left(2L_T(\boldsymbol{\theta}_{\boldsymbol{\xi}}) - c_T(\boldsymbol{\alpha}_T) n(\boldsymbol{\theta}_{\boldsymbol{\xi}}) \right) \right] \right\},
\tag{11}
$$

where $L_T(\boldsymbol{\theta_\xi})$ is the log-likelihood resulting for the selected model $\boldsymbol{\xi}$ at the parameter vector $\boldsymbol{\theta_\xi}$ for a sample $\mathbf{Y}_T$ of size T and $n(\boldsymbol{\theta_\xi})$ is the number of free parameters in the vector $\boldsymbol{\theta}$ associated with selection $\boldsymbol{\xi}$. The structure is similar to information criteria considered in the literature such as AIC with $c_T = 2$ (see Akaike, 1985), BIC with $c_T = \log T$ (see Schwarz, 1978), and HQ with $c_T = 2\log(\log T)$ (see Hannan and Quinn, 1979). But in contrast to those, the *Gets* selection process also ensures the congruence of a selected model, thus $\tilde{\boldsymbol{\xi}} \in \boldsymbol{\Xi}^c$ where $\boldsymbol{\Xi}^c$ is the subset of $\boldsymbol{\Xi}$ consisting of all congruent models.

The statistical properties of the proposed model selection procedure will me measured as the deviation of the selected model from the true model, $\left\| \tilde{\boldsymbol{\xi}} - \boldsymbol{\xi}^{\mathrm{DGP}} \right\|$, usually expressed in terms of size, power and the probability finding the truth. For the framework considered here, the implications of the selection in terms of the accuracy and precision of impulse responses and predictions is at least as important. But before we analyze the properties of the procedure proposed here within a realistic Monte Carlo example in §4, we investigate the critical issue of complexity.

System procedures have the disadvantage that the number of subset models of a K-dimensional VAR(p) with an unrestricted variance-covariance matrix is given by $2^{K^2 p}$ (without deterministic terms). Even for model selection procedures based on a single criterion such as the usual information criteria, a full search over all possible candidates is computationally unfeasible: in a K-dimensional VAR(p) without deterministic terms there are $K^2 p$ coefficients, any full search requires the estimation of a total of $2^{K^2 p}$ subset models. Already for a four-dimensional VAR(p), the computational costs are immense: a full search procedure has to check 65,536 subset models of order one, 4,294,967,296 for $p = 2$, 2.8×10^{14} for $p = 3$, 1.8×10^{19} for $p = 4$ and so on. The challenge for sequential simplification and testing procedure for the system, comparable to *PcGets* in single equations, is even greater. Such is the chance to miss the DGP in such a high-dimensional model universe even with the most advanced multi-path encompassing search algorithm. It is therefore imperative to the restrict the dimensionality of the model universe by decomposing the selection problem into manageable sub-tasks. Single equation procedure do so by partitioning the model space into K subspaces, $\boldsymbol{\Xi} = \boldsymbol{\Xi}_1 \times \cdots \times \boldsymbol{\Xi}_K$, with $\dim(\boldsymbol{\Xi}_k) = 2^{Kp}$. The critical question is whether there a loss in efficiency by analyzing the equations of a VAR once at a time using single-equation model selection algorithms rather than analyzing the VAR with a system procedure.

3.1 *Gets* reductions of reduced-form VAR models

We start by investigating *Gets* reductions of the reduced-form VAR(p) model defined in equation (1) in the case of a diagonal variance-covariance matrix, such that the equations of the system are unrelated to each other:

Proposition 1 (Reduction under conditional independence). *Suppose that in the reduced-form VAR defined in equation (1), the variance-covariance matrix $\boldsymbol{\Sigma}$ is diagonal, i.e. all $\sigma_{ij} = 0$ for $i \neq j$. Then, all possible reductions of the system can be efficiently estimated by OLS, and model-selection procedures can operate equation-by-equation without a loss in efficiency.*

Proof. Conditional independence of $\mathbf{y}_t$ conditional on its past $\mathbf{Y}_{t-1}$ allows the factorization of the probability density function of $\mathbf{y}_t$ in terms of its marginals:

$$f_y(\mathbf{y}_t | \mathbf{Y}_{t-1}) = \prod_{k=1}^{n} f_{y_k}(y_{kt} | \mathbf{Y}_{t-1}).$$

This implies that the log-likelihood function $L_T(\boldsymbol{\theta})$ can be separated with regard to the parameters of interest, $\boldsymbol{\theta}_k$, of each equation $k = 1, \ldots, K$ of the system which can be varied freely:

$$L_T(\boldsymbol{\theta}) = \sum_{k=1}^{K} \left\{ \sum_{t=1}^{T} \ln f_{y_k}(y_{kt}|\mathbf{Y}_{t-1}; \boldsymbol{\theta}_k) \right\} = \sum_{k=1}^{K} L_{kT}(\boldsymbol{\theta}_k).$$

Consequently, all possible reductions of the system can be (asymptotically) efficiently estimated by single equation methods (OLS under normality), and reduction procedures can be applied equation-by-equation without a loss of (asymptotic) efficiency. ∎

Proposition 1 states that the efficiency of single-equation model selection algorithms depends on the absence of instantaneous causality. In other words, if the variance-covariance matrix of the system is diagonal, *i.e.*, all $\sigma_{ij} = 0$ for $i \neq j$, the system can be analyzed as collection of single-equation models. Hence, single-equation reduction procedures can be applied under optimality conditions. In VAR models with instantaneous causality between the variables, the separability property of the log-likelihood function is lost: due to the contemporaneous correlation of variables in the system, the equations of the VAR are only seemingly unrelated to each other. Since eliminating a variable in one equation affects the others, single-equation model selection procedures are inefficient. This is directly related to the properties of the OLS estimation method, which is inefficient for subset VARs with non-diagonal variance-covariance matrices whereas full information maximum likelihood (FIML) and estimated generalized least squares (EGLS) are (asymptotically) efficient. For the efficient reduction of interdependent reduced-form VAR models, *Gets* algorithms have to be implemented as system procedures.

3.2 Efficiency of *Gets* single-equation reduction procedures for SVAR models

In the following proposition, we argue that the recursive structure of SVAR models as defined in (5) re-establishes conditions for the efficiency of single-equation *Gets* reduction procedures. The crucial point is that the researcher does not aim to reduce the reduced-form VAR, but the recursive SVAR:

Proposition 2 (Reduction under causal ordering). *Suppose that the GUM is a just-identified SVAR of the form defined by (5). Then, all possible reductions of the SVAR can be efficiently estimated by OLS, and model-selection procedures can operate equation-by-equation without a loss in efficiency.*

Proof. Causal ordering of $\mathbf{y}_t$ allows the factorization of the probability density function of $\mathbf{y}_t$ in terms of marginals and conditionals:

$$f_y(\mathbf{y}_t|\mathbf{Y}_{t-1}) = f_{y_1}(y_{1t}|\mathbf{Y}_{t-1}) \cdot f_{y_2|y_1}(y_{2t}|y_{1t}, \mathbf{Y}_{t-1}) \cdot \ldots \cdot f_{y_K|y_1,\ldots,y_{K-1}}(y_{Kt}|y_{1t}, \ldots, y_{K-1t}, \mathbf{Y}_{t-1}).$$

This implies that the log-likelihood function $L_T(\boldsymbol{\lambda})$ can be separated with regard to the parameters of interest, $\boldsymbol{\lambda}_k$, of each equation $k = 1, \ldots, K$ of the system:

$$\begin{aligned} L_T(\boldsymbol{\lambda}) &= \sum_{t=1}^{T} \ln f_{y_1}(y_{1t}|\mathbf{Y}_{t-1}, \boldsymbol{\lambda}_1) + \sum_{k=2}^{K} \left\{ \sum_{t=1}^{T} \ln f_{y_k|y_1,\ldots,y_{k-1}}(y_{kt}|y_{1t}, \ldots, y_{k-1t}, \mathbf{Y}_{t-1}, \boldsymbol{\lambda}_k) \right\} \\ &= \sum_{k=1}^{K} L_{kT}(\boldsymbol{\lambda}_k). \end{aligned}$$

In other words, $y_1, \ldots, y_{k-1}$ are weakly exogenous for the parameter vector $\boldsymbol{\lambda}_k = (\beta_{k1}, \ldots, \beta_{kk-1}, \gamma_{k1.1}, \ldots, \gamma_{kk.p}, \omega_k^2)'$ of the k-th equation of the SVAR, which can be varied freely. Thus, reductions procedures for the SVAR can be implemented as single-equation techniques. ∎

The great advantage of separability of the log-likelihood function for the recursive SVAR is that the model space generated by the just-identified SVAR which is of dimension $2^{K(Kp+(K-1)/2)}$, can be

searched in K subspaces of dimensions 2^{Kp} to $2^{Kp+(K-1)}$. In §4 this case is studied in a Monte Carlo experiment, where we use *PcGets* as the computer-automated single-equation *Gets* reduction procedure.

4 Monte Carlo results

Although the sequential nature of the proposed model-selection process and its combination of variable-selection and diagnostic testing has eluded most attempts at theoretical analysis, an evaluation of the properties of the model-selection process can be achieved by simulation. In the Monte Carlo (MC) experiment considered here, the properties of the *Gets* reduction of SVAR models are studied under the conditions of proposition 2, which allows the optimal use of single-equation reduction procedures. For the computer implementation of the proposed model selection procedure we use *PcGets*.

4.1 Design of the MC

The DGP is a three-dimensional Gaussian SVAR(1) model with the causal ordering $y_{1t} \rightarrow y_{2t} \rightarrow y_{3t}$:

$$
\begin{bmatrix} 1 & 0 & 0 \\ \beta_{21} & 1 & 0 \\ \beta_{31} & \beta_{32} & 1 \end{bmatrix} \begin{bmatrix} y_{1t} \\ y_{2t} \\ y_{3t} \end{bmatrix} = \begin{bmatrix} \gamma_{11} & 0 & \gamma_{13} \\ 0 & \gamma_{22} & \gamma_{23} \\ 0 & 0 & \gamma_{33} \end{bmatrix} \begin{bmatrix} y_{1t-1} \\ y_{2t-1} \\ y_{3t-1} \end{bmatrix} + \begin{bmatrix} \eta_{1t} \\ \eta_{2t} \\ \eta_{3t} \end{bmatrix} \tag{12}
$$

$$
\text{where } \boldsymbol{\eta}_t \sim \text{NID} \left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \omega_1^2 & 0 & 0 \\ 0 & \omega_2^2 & 0 \\ 0 & 0 & \omega_3^2 \end{bmatrix} \right).
$$

Its reduced-form is given by the VAR(1) model:

$$
\begin{bmatrix} y_{1t} \\ y_{2t} \\ y_{3t} \end{bmatrix} = \begin{bmatrix} \gamma_{11} & 0 & \gamma_{13} \\ -\beta_{21}\gamma_{11} & \gamma_{22} & \gamma_{23}-\beta_{21}\gamma_{13} \\ -\beta_{21}\gamma_{11} & -\beta_{32}\gamma_{22} & \gamma_{33}-\beta_{31}\gamma_{13}-\beta_{32}\gamma_{23} \end{bmatrix} \begin{bmatrix} y_{1t-1} \\ y_{2t-1} \\ y_{3t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \end{bmatrix} \tag{13}
$$

$$
\text{where } \boldsymbol{\varepsilon}_t \sim \text{NID} \left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \omega_1^2 & -\beta_{21}\omega_1^2 & -\beta_{31}\omega_1^2 \\ -\beta_{21}\omega_1^2 & \omega_2^2+\beta_{21}^2\omega_1^2 & -\beta_{32}\omega_2^2+\beta_{31}\beta_{21}\omega_1^2 \\ -\beta_{31}\omega_1^2 & -\beta_{32}\omega_2^2+\beta_{31}\beta_{21}\omega_1^2 & \omega_3^2+\beta_{31}^2\omega_1^2+\beta_{32}^2\omega_2^2 \end{bmatrix} \right).
$$

We can think of the following 'deep' structure associated with recursive structure of the SVAR:

$$
\begin{aligned}
\Delta y_{1t} &= -\alpha(y_{1t-1}-y_{3t-1})+\eta_1, \\
\Delta y_{2t} &= -\beta(y_{2t-1}-y_{3t-1}-y_{1t})+\eta_2, \\
\Delta y_{3t} &= \gamma(y_{2t}-y_{1t})-\rho y_{3t-1}+\eta_3,
\end{aligned}
$$

which for $\alpha = 0.4$, $\beta = 0.4$, $\rho = 0.4$, and $\gamma = 0.5$, results in the following parameterization of the SVAR in (12):

$$
\begin{bmatrix} 1 & 0 & 0 \\ 0.4 & 1 & 0 \\ 0.5 & -0.5 & 1 \end{bmatrix} \begin{bmatrix} y_{1t} \\ y_{2t} \\ y_{3t} \end{bmatrix} = \begin{bmatrix} 0.6 & 0 & 0.4 \\ 0 & 0.6 & 0.4 \\ 0 & 0 & 0.6 \end{bmatrix} \begin{bmatrix} y_{1t-1} \\ y_{2t-1} \\ y_{3t-1} \end{bmatrix} + \begin{bmatrix} \eta_{1t} \\ \eta_{2t} \\ \eta_{3t} \end{bmatrix} \tag{14}
$$

$$
\boldsymbol{\eta}_t \sim \text{NID} \left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right).
$$

Thus the reduced form of the SVAR in (14) is given by the VAR(1) model:

$$
\begin{bmatrix} y_{1t} \\ y_{2t} \\ y_{3t} \end{bmatrix} = \begin{bmatrix} 0.60 & 0 & 0.40 \\ -0.24 & 0.60 & 0.24 \\ -0.42 & 0.30 & 0.52 \end{bmatrix} \begin{bmatrix} y_{1t-1} \\ y_{2t-1} \\ y_{3t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \end{bmatrix} \tag{15}
$$

$$
\varepsilon_t \sim \text{NID} \left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1.00 & -0.40 & -0.50 \\ -0.40 & 1.16 & 0.78 \\ -0.50 & 0.78 & 1.74 \end{bmatrix} \right).
$$

With the eigenvalues of A_1 being 0.672 ± 0.351 and 0.376, the DGP is stationary with a zero mean. The sample size T is 100 and the number of replications M is 1000.

In the Monte Carlo experiment we intend to evaluate the properties of the proposed *Gets* selection procedure. For the computer-automation of the reduction process, we use the conservative strategy of Hendry and Krolzig (2001). In addition to the selection properties discussed in the next section, the analysis of the simulation experiment will focus on the accuracy and precision of the resulting impulse-responses.

4.2 Selection properties

Search commences from a just-identified SVAR(5) or unrestricted VAR(5) model. Although the just-identified structural VAR and the unrestricted reduced-form VAR are identical models, they generate different model spaces. Therefore a *Gets* reduction process commencing from one of these will al-most surely select a different reduced model than when starting with the other. An important difference between the selected VAR and SVAR arises from the fact that in subset VAR models the variance-covariance matrix remains unrestricted during the search, while in subset SVAR models the free elements of the matrix **B** will be subject to significance tests like any other coefficient of the system, resulting in a restricted variance-covariance matrix of its reduced form.

The selection properties of *Gets* reduction process proposed in this paper are reported in table 1 in terms of the ability of finding the DGP, the dominance of the selected model by the unknown DGP, as well as its size and power.

Given the characteristics of the DGP, the probability to find the DGP by *PcGets* is in between 78.8% and 86% for the SVAR, and 22.2% to 86% for the VAR. While naturally dependent on the design of DGP, these figures indicate that it is possible to retrieve the structure hidden in the data with a high probability. For the VAR, the success probabilities appear to be small, but they have to be compared to the probability of finding the DGP when starting the search from it, which is in between 33.3% and 100%. As the overall probability to find or miss the DGP is not very informative, we check by an encompassing test whether the deviation of the selected model from the DGP results in a sound model that, based on statistically criteria, could not have been improved by knowing the truth. As long as *PcGets* is able to find a model that is not dominated by the DGP itself, the reduction process has been a success. If the specific model is dominated by the DGP, the search algorithm has failed. Our results indicate that the risk of finding a model which is dominated by the DGP is relatively small (2.6% to 4.2%) when compared to the probability that the selected model dominates the truth, which is 3 to 15 times higher. Note that, by construction, the outcome of the *PcGets* search algorithm will always beat the unrestricted VAR(5) model.

Next we analyze the 'size' and 'power' of the proposed model-selection process, namely the probability of including variables that do not (do) enter the DGP in the selected model. The 'size' of *PcGets* (the average probability of selecting a nuisance regressor) is with 1.57% to 3.76% slightly higher than

Table 1 Selection properties.

Model	SVAR			VAR		
Equation	$y_{1,t}$	$y_{2,t}$	$y_{3,t}$	$y_{1,t}$	$y_{2,t}$	$y_{3,t}$
DGP found when commencing from it	1.000	0.982	0.999	1.000	0.333	0.386
DGP found by *PcGets*	0.860	0.788	0.843	0.860	0.222	0.282
Non-deletion probability	0.140	0.189	0.155	0.140	0.325	0.264
Non-selection probability	0.020	0.056	0.021	0.020	0.738	0.675
DGP dominated by *PcGets*	0.099	0.153	0.113	0.099	0.554	0.550
PcGets dominated by DGP	0.026	0.042	0.031	0.026	0.033	0.038
Size	0.0157	0.0232	0.0169	0.0157	0.0376	0.0305
Size (*reliability based*)	0.0116	0.0165	0.0129	0.0116	0.0315	0.0263
Power	0.9880	0.9770	0.9920	0.9880	0.7227	0.7420
Power (*reliability based*)	0.9874	0.9727	0.9912	0.9874	0.7062	0.7250

the nominal size of a t-test which is 1%. The size can be further improved by basing the selection decision on the reliability of the variables, which depends on the significance of the variables in two overlapping subsamples. In this case, the 'size' shrinks to 1.16% – 3.15%. The 'power' of *PcGets* (the average probability of selecting a DGP variable) is in between 99.2% and 72.27%. Overall, *PcGets* works more than satisfactorily despite the presence of collinearity among the regressors.

Having clarified the 'success' and 'failure' of *PcGets* in selecting a parsimonious representation of the structure found in the data, be proceed by checking how the selection properties of the proposed reduction strategy are translated into accuracy and precision of the impulse responses implied by the empirical model.

4.3 Impulse-response analysis

We now evaluate the properties the impulse-responses implied by the following models:

    (i) the true SVAR(1) defined in (14);
   (ii) the pseudo-true VAR(1) defined in (15);
  (iii) a just-identified SVAR(5) with intercept, which nests (14) and is identical to an unrestricted VAR(5) with intercept nesting (15);
  (iv) a *Gets* reduction of the just-identified SVAR(5) in (iii);
   (v) a *Gets* reduction of the VAR(5) in (v).

We are interested in the responses to structural and non-structural shock as laid out by the vector moving-average representations (3) and (9). The responses to the (normalized) structural shocks $\boldsymbol{\eta}_t^* = \boldsymbol{\Omega}^{-\frac{1}{2}}\boldsymbol{\eta}$ are given by:

$$\frac{\partial \mathbf{y}_{t+h}}{\partial \boldsymbol{\eta}_t^{*\prime}} = \boldsymbol{\Phi}_h = \mathbf{A}^h \mathbf{P} = (\mathbf{B}^{-1}\boldsymbol{\Gamma})^h \mathbf{B}^{-1}\boldsymbol{\Omega}^{\frac{1}{2}}. \tag{16}$$

The responses to the non-structural shocks $\boldsymbol{\varepsilon}_t$ result as:

$$\frac{\partial \mathbf{y}_{t+h}}{\partial \boldsymbol{\varepsilon}_t'} = \boldsymbol{\Psi}_h = \mathbf{A}^h = (\mathbf{B}^{-1}\boldsymbol{\Gamma})^h. \tag{17}$$

The impulse-responses functions implied by the models (i) to (v) are illustrated in figure 1, without loosing generality, for the response of the third variable y_{3t+h} to shocks in the first variable y_{1t} of the system. The graph on the left displays the responses of y_{3t+h} to the structural shock η_{1t}. Plotted are their mean response (over the $M = 1000$ replications) and their 90% confidence intervals as implied
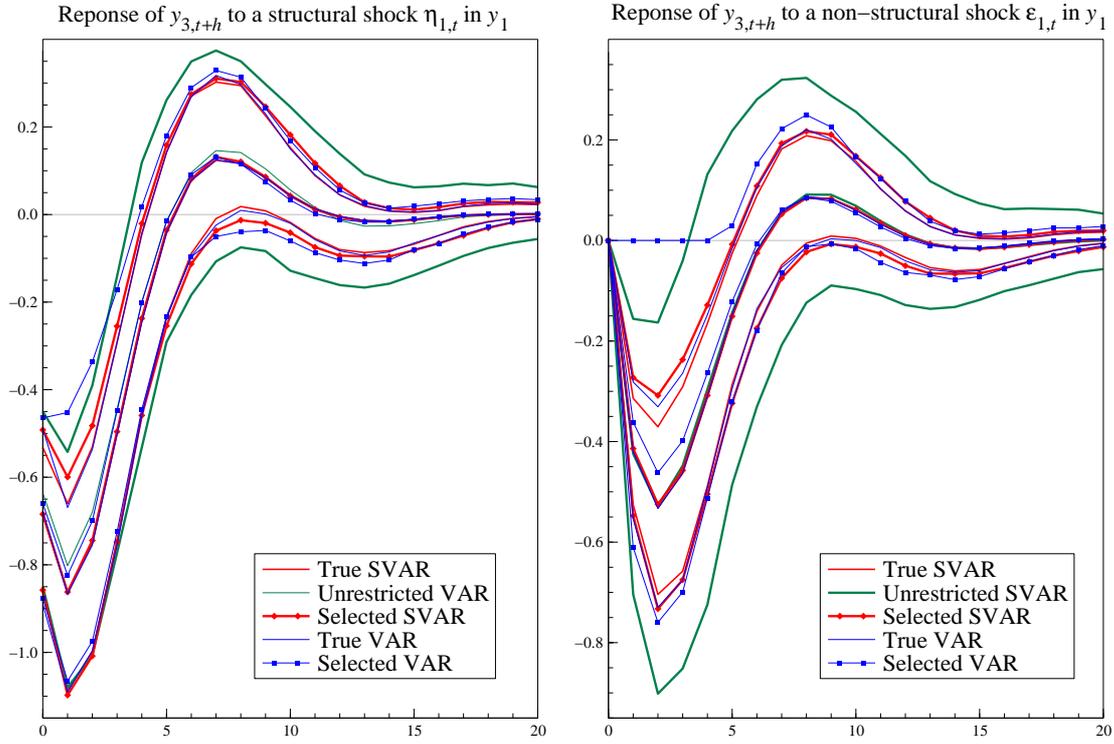
Figure 1 Impulse-response function: mean and 90% confidence intervals.

by the 0.05 and 0.95 quantiles. The graph on the right shows the responses to the non-structural shock ε_{1t}. Note that the 90% confidence intervals for impulse-responses of the true and unrestricted SVAR as well as reduced-form VAR only reflect estimation uncertainty, while the confidence intervals for impulse-responses of the models selected by *PcGets* also account for the specification uncertainty.

Except for the subset VAR models, the mean impulse responses plotted in figure 1 are all close to each other, indicating that consistent estimates of the impulse responses can not only be obtained from unrestricted VAR models, but also from the selected SVAR model. Although the selected SVAR will always differ from the DGP with a positive probability, the squared biases (and confidence intervals) of its impulse responses are not much greater than for the true SVAR, when its structure but not its parameters are known and necessitates estimation. The bias problem caused by single-equation based reductions of the reduced-form VAR is nicely illustrated in figure 1: Some of the dynamic multipliers are falsely shrunk to zero as statistically insignificant autoregressive parameters are eliminated from the model

Figure 2 exhibits the squared bias of all structural and non-structural impulse responses. The bias can be defined as:

$$\text{Bias}\left[\phi_{ij,h}\right] = \bar{\phi}_{ij,h} - \phi_{ij,h} \tag{18}$$

where $\phi_{ij,h}$ denotes the theoretical impulse response as implied by the model in equation (14) and $\bar{\phi}_{ij,h}$ is the mean impulse response of one of the empirical models (i) to (v):

$$\bar{\phi}_{ij,h} = \frac{1}{M} \sum_{m=1}^{M} \tilde{\phi}_{ij,h}^{(m)}. \tag{19}$$

In a completely analogous fashion, the squared bias of the non-structural impulse-responses can be calculated. While for the true SVAR, the pseudo-true reduced-form VAR and the selected SVAR, the
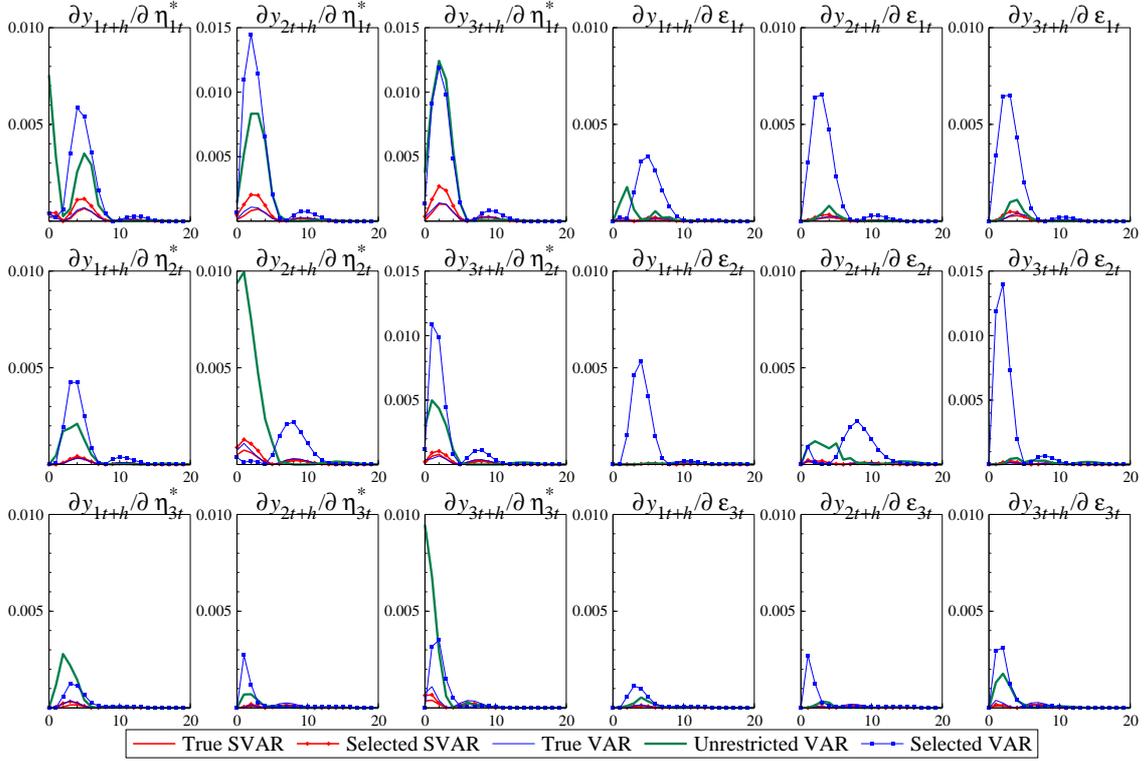
Figure 2 Squared bias of structural and non-structural impulse-responses.

accuracy of the non-structural impulse-responses is as good as that of the structural impulse responses, the unrestricted VAR looses accuracy when structural impulse responses are considered. This is because the Choleski decomposition fails to detect the structure in the variance-covariance matrix resulting from the zero parameters in the $\mathbf{B}$ matrix.

The uncertainty associated with the estimated impulse-responses can be measured by their variance:

$$\mathsf{Var}\left[\phi_{ij,h}\right] = \frac{1}{M} \sum_{m=1}^{M} \left(\tilde{\phi}_{ij,h}^{(m)} - \bar{\phi}_{ij,h}\right)^2, \tag{20}$$

which for the Gaussian framework analyzed here, allows the construction of confidence bands in the usual way. For the non-structural impulse-responses, we analyze the elements of $\mathbf{\Psi}_h$ instead of $\mathbf{\Phi_h}$. Figure 3 plots the variances of estimated structural and non-structural impulse-responses for the modelling strategies (i) to (v). Limiting the number of parameters by model reduction clearly helps to reduce estimation uncertainty. Consequently, the selected models produce more precisely estimated impulse responses than the full VAR. As seen in figure 1, the confidence bands for the impulse responses of the selected SVAR are much closer to those of the true SVAR than the ones of the just-identified SVAR. For this purpose, also single-equation based reductions of the reduced-form VAR are beneficial.

Finally, the overall error in estimating the impulse responses can be quantified by the mean square error:

$$\mathsf{MSE}\left[\phi_{ij,h}\right] = \frac{1}{M} \sum_{m=1}^{M} \left(\tilde{\phi}_{ij,h}^{(m)} - \phi_{ij,h}\right)^2 = \mathsf{Bias}^2\left[\phi_{ij,h}\right] + \mathsf{Var}\left[\phi_{ij,h}\right] \tag{21}$$

for the response of $y_{i,t+h}$ to a unit impulse to the structural shock term $\eta_{j,t}$ and accordingly by $\mathsf{MSE}\left[\psi_{ij,h}\right]$ for the response to a non-structural shock in equation j, $\varepsilon_{j,t}$. The results are displayed
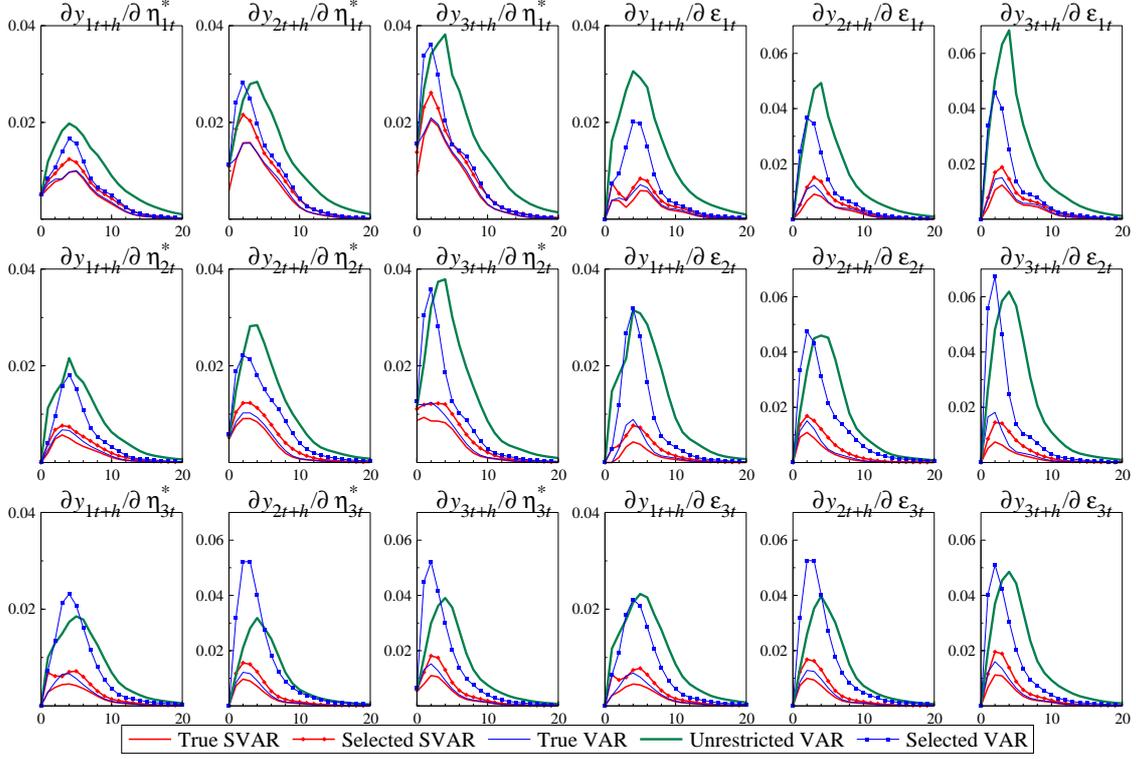
Figure 3 Variance of structural and non-structural impulse-responses.

in figure 4 with the structural impulse responses on the left and non-structural impulse responses on the right.

As the MSE summarizes the effects of bias and variance, we get a clear picture of the relative precision and accuracy of the impulse response functions generated by modelling strategies (i) to (v). The true SVAR naturally delivers the best results throughout, followed by the pseudo-true VAR, and then the SVAR selected by *PcGets*. These three models consistently dominate the unrestricted VAR as well as the selected VAR. For $h < 8$, the unrestricted VAR and the equation-by-equation reduced VAR are close competitors, but for longer horizons, the MSEs of the subset VAR decline faster. Since the DGP and the estimated models are stable, the theoretical as well as the empirical impulse responses fade out with increasing distance between action and reaction. The MSEs converge to zero with increasing h producing the hump shaped plots in figure 4.

The precision and accuracy of the model selection strategies are summarized in table 2 by calculating simple averages of the measures in (18) to (21) such as

$$\mathsf{MSE}[\phi] = \frac{1}{K^2 H} \sum_{i=1}^{K} \sum_{j=1}^{K} \sum_{h=1}^{H} \mathsf{MSE}\left[\phi_{ij,h}\right], \tag{22}$$

for the non-structural impulse-responses and, analogously, $\mathsf{MSE}[\psi]$ for the non-structural impulse-responses. For the sake of convenience, the results have been normalized to one for the unrestricted VAR (just identified SVAR).

Table 2 provides a clear ranking of the modelling strategies: The true SVAR delivers the best results, followed by the pseudo-true VAR and the SVAR selected by proposed *Gets* model reduction procedure. The loss in accuracy due to 'reconstruction' of the unknown structure is very small when compared to the loss in accuracy when working with an unrestricted VAR (just-identified SVAR). Reducing the reduced-
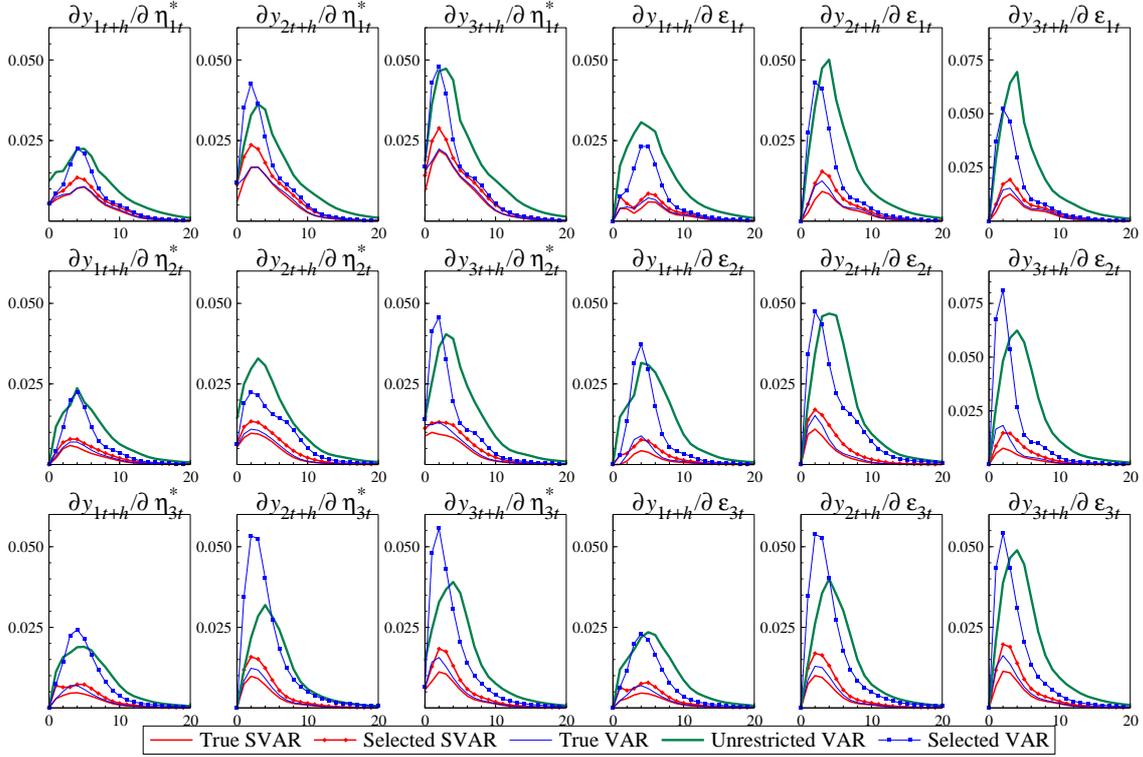
Figure 4 Mean square error of structural and non-structural impulse-responses.

Table 2 Precision and accuracy of impulse response functions.

		structural impulse responses			non-structural impulse responses		
		$Bias^2$	Variance	MSE	$Bias^2$	Variance	MSE
true SVAR	(i)	0.1244	0.3106	0.2945	0.2353	0.1489	0.1498
pseudo-true VAR	(ii)	0.1536	0.3559	0.3384	0.3204	0.2013	0.2025
selected SVAR	(iv)	0.1992	0.4465	0.4252	0.3112	0.2604	0.2610
selected VAR	(v)	0.9912	0.7916	0.8088	5.2855	0.6631	0.7105
unrestricted (S)VAR	(iii)	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

form VAR with single-equation methods when the system is interdependent can cause a substantial bias; but it avoids the inflation in variance, which affects the full VAR so badly.

Altogether, our MC results highlight the dangers of (i) using impulse response analysis for unrestricted, richly parameterized VAR models and (ii) employing single-equation reduction procedures for reduced-form VAR models when the conditions stated in proposition 1 are not met. While our Monte Carlo results suggest that the case for using *Gets* reductions of SVAR models for impulse-response analysis is a strong one, it is still worth emphasizing two major limitations of the analysis presented here: First, a larger variety of DGPs has to be considered before we can conclude that the features found here are indeed systematic properties of proposed *Gets* procedure for the reduction of SVAR models. Secondly, it would be interesting to compare the *Gets* reduction procedure proposed in this paper with other model selection strategies discussed in the literature. For VAR models satisfying the conditional independence condition of proposition 1, Brüggemann *et al.* (2002) compared alternative computerized model-selection strategies and found a clear advantage of the *PcGets* algorithm in forecast comparisons. Similar competitions in the context of the SVAR modeling framework considered here are highly desirable.

5 Empirical illustration

To illustrate the proposed *Gets* procedures for SVARs, we will now use *PcGets* to analyze a small macro-econometric model for the US. The model is the monetary system introduced by Christiano *et al.* (1996) consisting of the log of real GDP, gdp, the log of the GDP deflator, p, the log of a commodity price index, $pcom$, the Fed funds rate, ff, the negative log of unborrowed reserves, $nbrd$, the log of total reserves, tr and the log of M1, $m1$.
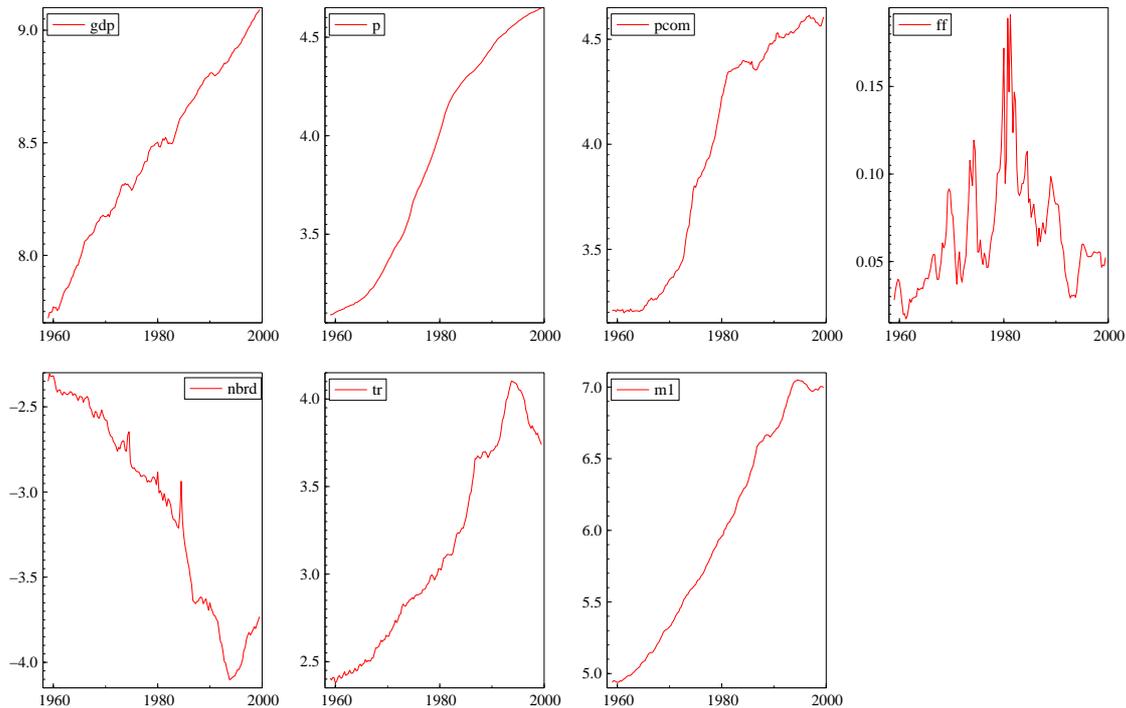
Figure 5   The extended Christiano *et al.* (1996) data set.

The data are in levels and are plotted in Figure 5. The graphs show trending behavior in all variables. This could potentially cause a problem as, to date, *PcGets* conducts all inferences as I(0). The implementation of cointegration tests and appropriate transformations would be useful (see the discussion in §6). But most selection tests will in fact be valid even when the data are I(1), given the results in, *inter alia*, Sims, Stock and Watson (1990). Only t- or F-tests for an effect that corresponds to a unit root require non-standard critical values. Similarly, Wooldridge (1999) shows that diagnostic tests for the unrestricted model remain valid even for integrated time.

Christiano *et al.* (1996) considered an unrestricted reduced-form VAR(4) of the variables. Extending the data set to the period 1960 (i) to 1999 (iii), the VAR requires the estimation of 203 coefficients. Only few turn out to be significant at a 5% significance level (see table 3), calling for reductions of the VAR. Using the liberal strategy of *PcGets*, the reduced-form VAR can be simplified to the subset VAR summarized in table 4. The results are in line with the earlier studies of Brüggemann and Lütkepohl (2000) and Krolzig (2001).

There are two major problems with this approach: First, the misspecification of all equations of the VAR except the price level p (see table 5). There is a structural break in the middle of the sample period, which is presumably related to the Volcker deflation and affects the equations for commodity

Table 3 Unrestricted reduced-form VAR.

variable	ν	gdp				p				pcom				ff				nbrd				tr				m1			
lag	0	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
gdp	.	+	.	.	.	.	.	.	.	.	.	.	.	−	.	.	.	.	.	.	.	.	.	.	.	.	+	.	.
p	−	.	.	.	.	+	−	.	−	+	.	.	.	−	.	.	.	+	−	.	.	.	.	+	.	.	.	.	.
pcom	.	.	.	.	.	+	−	.	.	+	.	.	−	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
ff	.	+	.	.	.	.	.	−	+	+	.	.	.	+	.	+	.	.	.	.	.	.	.	.	.	.	+	−	.
nbrd	.	+	.	.	−	.	.	.	.	+	.	.	.	.	.	.	.	+	−	.	.	−	.	.	.	.	.	.	.
tr	.	−	.	.	+	.	.	.	.	.	.	.	.	+	.	.	.	.	.	.	.	+	.	+	+	.	.	−	.
m1	−	−	.	.	+	.	.	.	+	−	+	−	.	−	+	−	.	.	.	.	.	+	.	.	.	+	.	.	−

Legend: 0 Coefficient is set to zero.   + Coefficient is positive and significant at the 5% level.
       . Coefficient is insignificant at the 5% level.   − Coefficient is negative and significant at the 5% level.

Table 4 Subset reduced-form VAR.

variable	ν	gdp				p				pcom				ff				nbrd				tr				m1			
lag	0	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
gdp	0	+	0	0	0	+	0	0	0	0	0	0	0	−	0	0	0	0	0	0	0	−	0	0	0	0	0	0	0
p	−	0	0	+	0	+	0	0	−	+	0	−	0	0	0	0	0	0	0	0	0	+	0	−	+	0	0	0	−
pcom	0	0	0	0	0	+	−	0	0	+	−	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ff	0	+	0	−	0	0	0	−	0	+	0	0	−	+	0	+	0	−	0	0	0	−	0	0	0	+	0	0	−
nbrd	0	+	0	−	0	0	0	−	0	+	0	0	0	0	+	0	+	−	0	0	−	−	0	0	0	+	0	0	0
tr	0	−	0	0	+	0	0	0	+	−	0	0	0	0	0	0	−	0	+	0	+	+	0	0	+	+	−	0	−
m1	−	0	0	+	0	0	0	+	−	0	0	0	−	+	0	0	0	0	0	+	0	0	0	+	0	0	0	0	−

Legend: 0 Coefficient is set to zero.   + Coefficient is positive and significant at the 5% level.
       . Coefficient is insignificant at the 5% level.   − Coefficient is negative and significant at the 5% level.

prices pcom, Fed funds rate, ff, unborrowed reserves, nbrd, total reserves, tr, and M1. The residuals of gdp, pcom, ff and nbrd are non-Gaussian. There is some autocorrelation left in the case of M1, heteroscedasticity and ARCH effects can be found in the pcom and ff equations. Secondly, there is contemporaneous correlation among the residuals of the system. This violation of the condition for the efficiency of single-equation reduction procedures as stated in proposition 1 sheds some doubt on the properties of the reduction process, despite the fact that the over-identifying restrictions are not rejected by the LR test: $\chi^2(143) = 150.44[0.3186]$.

Table 5 Misspecification tests: Unrestricted reduced-form VAR.

	gdp	p	pcom	ff	nbrd	tr	m1
Chow(1979:4)	0.9264	0.5936	**0.0076**	**0.0000**	**0.0000**	**0.0025**	**0.0074**
Chow(1995:4)	0.5470	0.9800	0.8844	0.9833	0.9971	0.0479	0.4010
normality test	**0.0078**	0.2585	**0.0000**	**0.0000**	**0.0000**	0.8698	0.2836
AR 1-4 test	0.0492	0.0424	0.0414	0.1169	0.0107	0.5876	**0.0006**
ARCH 1-4 test	0.9620	0.5889	**0.0001**	**0.0000**	0.2996	0.5952	0.0656
hetero test	0.9710	0.6477	0.0414	**0.0000**	0.2898	0.3695	0.0763

Reported are the marginal rejection probabilities. Tests in bold are significant at 1%.

Given the strong indication of instantaneous causality, it seems appropriate to adopt the causal ordering imposed by Christiano *et al.* (1996),

$$\text{gdp}_t \rightarrow \text{p}_t \rightarrow \text{pcom}_t \rightarrow \text{ff}_t \rightarrow \text{nbrd}_t \rightarrow \text{tr}_t \rightarrow \text{m1}_t,$$

in their analysis of the effects of monetary policy shocks using orthogonalized impulse responses. The causal ordering is now used to set up the just-identified recursive SVAR. The involved sequential conditioning will admit the application of single-equation *Gets* reduction procedures.

Table 6   Just-identified structural VAR.

variable	ν	gdp 1	2	3	4	p 1	2	3	4	pcom 1	2	3	4	ff 1	2	3	4	nbrd 1	2	3	4	tr 1	2	3	4	m1 1	2	3	4	gdp 0	p 0	pcom 0	ff 0	nbrd 0	tr 0
gdp	.	+	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	−	.	.	.	.	.	.	.	.	.	.	+	.	.	.	.	.	.
p	−	.	.	.	.	+	−	.	−	+	.	.	.	.	.	.	.	+	−	.	.	.	−	+	.	.	.	.	.	.	.	.	.	.	.
pcom	.	.	.	.	.	.	.	+	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	+	.	.
ff	.	.	.	.	.	.	−	+	.	.	.	.	+	.	+	.	.	.	.	.	.	.	.	.	.	.	.	+	−	+	.	+	.	.	.
nbrd	.	.	.	.	−	.	.	.	.	.	.	.	.	.	−	.	.	.	+	−	.	.	.	−	.	.	.	.	.	.	.	+	+	.	.
tr	.	.	.	.	.	.	.	.	.	.	.	.	−	.	−	+	.	.	+	.	.	+	+	.	.	+	+	.	.	.	.	.	.	+	−
m1	−	−	+	.	.	.	.	−	+	−	.	.	.	+	−	.	.	.	.	.	−	.	.	.	+	.	.	.	+	+	.	−	.	+	.

Legend: 0 — Coefficient is set to zero. + — Coefficient is positive and significant at the 5% level.
. — Coefficient is insignificant at the 5% level. − — Coefficient is negative and significant at the 5% level.

Table 7   Over-identified structural VAR.

variable	ν	gdp 1	2	3	4	p 1	2	3	4	pcom 1	2	3	4	ff 1	2	3	4	nbrd 1	2	3	4	tr 1	2	3	4	m1 1	2	3	4	gdp 0	p 0	pcom 0	ff 0	nbrd 0	tr 0
gdp	0	+	0	0	0	+	0	0	0	0	0	0	0	0	−	0	0	0	0	0	0	0	−	0	0	0	0	0	0						
p	−	0	0	+	0	+	−	+	−	+	0	−	0	0	0	0	0	0	0	0	0	+	0	−	0	0	0	0	0	0					
pcom	0	0	0	0	0	−	0	0	0	+	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	+					
ff	0	0	0	−	0	0	0	−	0	0	−	0	0	+	0	+	0	0	0	0	0	−	0	0	0	+	0	0	0	+	0	+			
nbrd	0	0	0	0	0	0	0	0	0	0	0	0	0	−	0	0	0	+	−	0	0	0	−	0	0	−	+	0	0	0	0	0	+		
tr	0	0	0	0	0	−	+	0	0	0	0	0	0	−	+	0	0	+	0	0	0	+	0	0	0	+	0	−	0	0	0	0	+	−	
m1	−	0	0	0	0	0	0	0	+	0	0	0	0	0	+	0	0	0	0	0	0	−	0	0	0	+	0	0	−	+	0	0	−	0	+

Legend: 0 — Coefficient is set to zero. + — Coefficient is positive and significant at the 5% level.
. — Coefficient is insignificant at the 5% level. − — Coefficient is negative and significant at the 5% level.

Table 6 reports the properties of the exactly identified SVAR which is observationally equivalent to the unrestricted VAR in table 3.[1] Interestingly, conditioning removes the structural break in the commodity price, $pcom_t$, and the money demand equation, $m1_t$ indicating the presence of cobreaking among the variables of the system. But, overall, table 8 shows that the (just identified) SVAR is also misspecified stressing the necessity of continued research on the formulation of a congruent representation of the data.

Table 8   Misspecification tests: Just-identified structural VAR.

	gdp	p	pcom	ff	nbrd	tr	m1
Chow(1979:4)	0.9264	0.5634	0.0450	**0.0000**	**0.0000**	**0.0076**	0.3329
Chow(1995:4)	0.5470	0.9730	0.6778	0.9797	0.9977	0.0698	0.6214
normality test	**0.0078**	0.2899	**0.0001**	**0.0000**	**0.0000**	0.1859	0.6647
AR 1-4 test	0.0492	0.0505	0.0644	0.0698	0.2272	0.0382	0.2558
ARCH 1-4 test	0.9620	0.6709	**0.0001**	**0.0000**	0.4869	0.4715	0.3567
hetero test	0.9710	0.7751	0.1622	**0.0000**	0.5533	**0.0008**	0.2833

Reported are the marginal rejection probabilities. Tests in bold are significant at 1%.

The properties of the selected SVAR are summarized in table 7. The contemporaneous relationships are found to be:

$$pcom_t(\; p_t \;),\; ff_t(\; gdp_t \;,\; pcom_t \;),\; nbrd_t(\; ff_t \;),\; tr_t(\; ff_t \;,\; nbrd_t \;),\; m1_t(\; gdp_t \;,\; ff_t \;,\; tr_t \;).$$
$$\quad\;\; + \qquad\qquad + \qquad\;\; + \qquad\qquad\quad + \qquad\qquad + \quad\;\; - \qquad\qquad\quad + \quad\;\; - \quad\;\; +$$

The restrictions imposed by the reduction process can not be rejected by an LR test of the overidentifying restrictions: $\chi^2(151) = 154.51\; [0.4056]$.

[1]A complete listing of the estimated equations of the models can be send by the author on request.
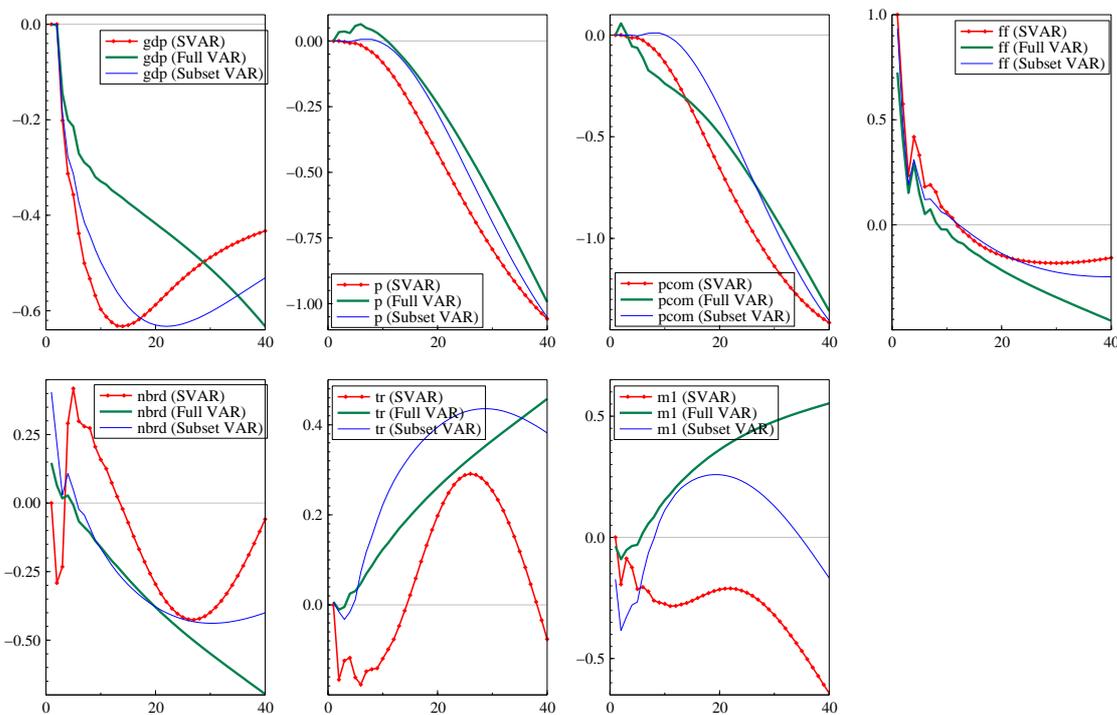
Figure 6 Response to a monetary policy shock.

Figure 6 shows the response of all system variables to a monetary policy shock, defined as a unit impulse in the Fed funds rate equation. Plotted are the responses derived from the (i) selected structural VAR, (ii) the selected reduced-form VAR and (iii) the full VAR as considered by Christiano *et al.* (1996). In case of a reduced-form VAR, the impulse-responses have been orthogonalized using the FIML (OLS) estimate of the variance-covariance matrix of the subset (full) VAR.

All three parametric models considered predict that an increase in the fed funds rate will cause a persistent drop in the level of GDP. However, while GDP is steadily falling in the unrestricted VAR model, the contraction in GDP bottoms out after 14 quarters in the over-identified SVAR and after 22 quarters in the subset VAR. The responses of the aggregate and commodity price indices show, after some delay, a smooth decline for the VAR and its reductions. The pattern of the own response of the fed funds rate over time is again very similar in the selected SVAR, the subset VAR and the full VAR. However, there are some striking differences in the impulse response functions of the selected models with regard to the monetary aggregates when compared to the impulse responses of the unrestricted system: While the full VAR predicts a secular decline or increase in the aggregates, the reaction of the negative log of unborrowed reserves and total reserves are mean-reverting in the structural VAR. Interestingly, the SVAR and the full VAR differ greatly regarding the responses in M1. Since we have shown earlier, that the reduced-form equation for M1 is misspecified (see table 5), whereas the structural equation for M1 was found to be congruent and stable over time (see table 8), we are inclined to put greater trust in the impulse response function derived from the selected SVAR. Finally, it is worth noting that since fewer parameters have to be estimated, the responses are estimated more precisely, making them a useful framework for tests for Granger causality etc.

6 Directions for further research

The *Gets* approach for the reduction of SVAR models presented in this paper has three important limitations: (i) the assumption that the causal ordering of the variable is known *a priori*, (ii) the stationarity of the data-generating process, and (iii) the recursive structure of simultaneous relations in the SVAR. In the following we briefly outline three directions for further research to overcome these limitations in a more general setting allowing for the applicability of the reduction procedure discussed here.

6.1 Selecting the causal order

The approach to the reduction of structural VARs proposed in this paper has been based on the assumption that the causal ordering of the variable is known *a priori*. In practice, the true causal order of the variables is usually unknown. While insight from the modeling context, particularly economic theory, can be extremely fruitful, there may be no unique ordering of the variables available. This raises the question whether sample evidence can be exploited for the selection of the causal order or, more precisely, the order of sequential conditioning.

While SVAR models which impose over-identifying restrictions can be tested, this is not the case for just-identifying restrictions. There is a rich literature on the determination of the design of the matrix $\mathbf{B}$ in (5) from the data by using the partial autocorrelations implied the estimated variance covariance matrix $\tilde{\Sigma}$ of the reduced-form VAR (see, *inter alia*, Swanson and Granger, 1997, Reale and Tunnicliffe Wilson, 2002, and Selva and Hoover, 2002). These approaches have in common a two-stage approach: (i) the determination of the causal order based on the estimated covariance matrix of the unrestricted system, and then (ii) the reduction of the VAR dynamics conditional on (i).

In contrast, we suggest the reduction of the fully-identified SVAR for all possible $K!$ causal orderings, and then the selection of the dominant model with the help of a consistent information criterion. If the K-dimensional vector of random variables $\mathbf{y}_t$ is generated by a structural VAR in (1) overidentified by restrictions on $\mathbf{\Gamma}$, then the reduced-form VAR and all other factorizations of the joint density should not be parsimoniously encompassing. A testing strategy would be preferable, but complications arise from the fact that even over-identified SVARs with different causal orders can have the same reduced form. Consider, for example, the following SVARs:

$$\mathbf{B}\mathbf{y}_t = \mathbf{\Gamma}\mathbf{y}_{t-1} + \boldsymbol{\eta}_t, \qquad \mathsf{E}[\boldsymbol{\eta}_t\boldsymbol{\eta}_t'] = \mathbf{\Omega} \tag{23}$$

$$\mathbf{B}^*\mathbf{y}_t = \mathbf{\Gamma}^*\mathbf{y}_{t-1} + \boldsymbol{\eta}_t^*, \quad \mathsf{E}[\boldsymbol{\eta}_t^*\boldsymbol{\eta}_t^{*\prime}] = \mathbf{\Omega}^*. \tag{24}$$

If $\mathbf{B}^{-1}\mathbf{\Gamma} = \mathbf{B}^{*-1}\mathbf{\Gamma}^* = \mathbf{A}$ and $\mathbf{B}^{-1}\mathbf{\Omega}\mathbf{B}^{-1\prime} = \mathbf{B}^{*-1}\mathbf{\Omega}^*\mathbf{B}^{*-1\prime} = \mathbf{\Sigma}$, then both models have the following reduced-form representation:

$$\mathbf{y}_t = \mathbf{A}\mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t, \quad \text{with } \mathsf{E}[\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'] = \mathbf{\Sigma}. \tag{25}$$

Hence, the SVARs (23) and (24) are observationally equivalent (see Hendry and Mizon, 2000, for more details).

6.2 *Gets* system reduction procedures for cointegrated VAR models

So far we assumed that the VAR is stable. Most economic data show stochastic trends, so that we should allow for integrated and possibly cointegrated processes. Since cointegration is a common feature of the variables in the system, and imposes nonlinear cross-equation restrictions on the parameters of the VAR, only system reduction procedures can be efficient. In the following, we outline a *Gets* system procedure for the reduction of cointegrated VAR models.

The reduced-form VAR(p) model in (1) can always be represented as a vector equilibrium correction model (VECM) of order $p - 1$:

$$\Delta \mathbf{y}_t = \boldsymbol{\nu} + \boldsymbol{\Pi} \mathbf{y}_{t-1} + \sum_{i=1}^{p-1} \boldsymbol{\Upsilon}_i \Delta \mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t, \quad \boldsymbol{\varepsilon}_t \sim \mathsf{NID}(\mathbf{0}, \boldsymbol{\Sigma}). \tag{26}$$

If $\boldsymbol{\Pi}$ has full rank, the variables $\mathbf{y}_t$ are $I(0)$, and without further restriction, the VECM is just-identified. An identification issue arises when $0 < \mathrm{rank}(\boldsymbol{\Pi}) = r < K$, in which case $\mathbf{y}_t \sim I(1)$ and there exist r cointegration vectors $\boldsymbol{\beta}' \mathbf{y}_t \sim I(0)$. Various methods for the cointegration analysis of multiple time series have been proposed in the literature. As the *Gets* procedures discussed here are likelihood based, Johansen's concentrated-likelihood-function approach (see Johansen, 1995) is the natural choice and suggests a three-stage reduction approach:

(1) *Reductions of the system.* Johansen's reduced rank procedures are based on reduced-form VAR without any equation-specific restrictions. Therefore a reduction procedure for the system

$$\Delta \mathbf{y}_t = \boldsymbol{\nu} + \boldsymbol{\Pi} \mathbf{y}_{t-1} + \sum_{j=1}^{K} \sum_{i=1}^{p} \boldsymbol{\Upsilon}_{ki} \Delta \mathbf{y}_{j,t-i} + \boldsymbol{\varepsilon}_t, \tag{27}$$

precedes the cointegration analysis. This step should involve a straightforward, multivariate generalization of the *PcGets* algorithm for single equation models aiming to simplify equation (27) by sequential tests for significance of the $\boldsymbol{\Upsilon}_{ki}$ parameter vectors involving block presearch reductions (including a lag selection procedure to determine the order of the VAR as discussed in Hendry and Krolzig, 2003) and a multiple-path encompassing search. The system procedure involves joint reductions of the system, *i.e.* the analysis of cross-equation restrictions, whose acceptance would exclude a regressor from all equations of the system (instead of individual equations as so done far),

(2) *Johansen cointegration tests* and *identification of the cointegration vectors*. The Johansen procedure for empirically determining the cointegration rank r is then applied to the reduction of system (27). It produces unique estimates of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ as a result of requiring $\boldsymbol{\beta}$ to be orthogonal and normalized. This estimate provides a value for the unrestricted log-likelihood function to be compared to the value for the log-likelihood function under overidentifying restrictions for $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ which have an economic interpretation. Omtzig (2002) has proposed an algorithm for automatic selection of cointegration vectors.

(3) *Reductions of the VECM*

   (i) *Reduced-form representation.* Given the outcome of the cointegration analysis (the cointegration rank r, cointegration matrix $\boldsymbol{\beta}$ and the structure of $\boldsymbol{\alpha}$), the analysis then focuses on reductions of the short-run dynamics $\boldsymbol{\Upsilon}_i$ in the corresponding (stationary) reduced-form VECM($p - 1$),

$$\Delta \mathbf{y}_t = \boldsymbol{\nu} + \boldsymbol{\alpha}_r \left(\boldsymbol{\beta}_r' \mathbf{y}_{t-1} \right) + \sum_{i=1}^{p-1} \boldsymbol{\Upsilon}_i \Delta \mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t, \quad \boldsymbol{\varepsilon}_t \sim \mathsf{NID}(\mathbf{0}, \boldsymbol{\Sigma}). \tag{28}$$

   Here, reduction paths for the system and the individual equations can be considered:

   (a) System searches for the coefficient with the lowest t-value of the system:

$$(k^*, j^*, i^*) := \arg \min_{k=1,\ldots,K} \min_{j=1,\ldots,K} \min_{i=1,\ldots,p} \mathsf{t}_{kj,i}^2.$$

(b) Single-equation searches for the coefficient with the lowest t-value of the k-th equation:

$$(j^*, i^*) := \arg \min_{j=1,\dots,K} \min_{i=1,\dots,p} \mathsf{t}^2_{kj,i} \text{ for } k = 1, \dots, K.$$

If the coefficient a_{kj^*,i^*} of regressor $y_{j,t-i}$ in equation k is insignificant, the coefficient is restricted to zero and the equation is re-estimated by OLS. Alternatively, EGLS could be used whereby the variance-covariance matrix is taken from the reduced, but otherwise unrestricted system.

Proposition 1 still applies: the single-equation analysis will only be efficient if the variance-covariance matrix is diagonal.

(ii) *Structural representation.* For a given recursive contemporaneous structure, we might be instead interested in the coefficients collected to the $\boldsymbol{\Upsilon}^*_i$ matrices in the structural VECM

$$\mathbf{B}\Delta\mathbf{y}_t = \boldsymbol{\delta} + \boldsymbol{\alpha}^*_r \left(\boldsymbol{\beta}'_r \mathbf{y}_{t-1}\right) + \sum_{i=1}^{p-1} \boldsymbol{\Upsilon}^*_i \Delta\mathbf{y}_{t-1} + \boldsymbol{\eta}_t, \quad \boldsymbol{\eta}_t \sim \mathsf{NID}(\mathbf{0}, \boldsymbol{\Omega}). \tag{29}$$

where $\mathbf{B}$ is a lower-triangular matrix and $\boldsymbol{\Omega}$ is a diagonal matrix. Identification of (29) follows from the restrictions formulated in (8) and the equivalence with (27) is given by $\boldsymbol{\Sigma} = \mathbf{B}^{-1}\boldsymbol{\Omega}\mathbf{B}^{-1'}$, $\boldsymbol{\alpha}_r = \mathbf{B}^{-1}\boldsymbol{\alpha}^*_r$, $\boldsymbol{\Upsilon}^*_i = \mathbf{B}^{-1}\boldsymbol{\Upsilon}_i$ for $i = 1, \dots, p-1$, and $\boldsymbol{\nu} = \mathbf{B}^{-1}\boldsymbol{\delta}$. Equation (29) becomes the new GUM to which the single-equation *Gets* reduction procedure proposed in this paper can be applied.

Alternatively, if the cointegration vectors are delivered by economic theory, the first two steps are made redundant and the *Gets* reduction approach serves to simplify the short-run dynamics only.

6.3 *Gets* reduction procedures for simultaneous equation models

The recursive contemporaneous structure and closeness of the SVAR in (5) can cause problems if the underlying economic theory predicts the interdependence of the endogenous variables in the temporary macroeconomic equilibrium. Fortunately it is possible to relax the assumption of a causal ordering of the variables made in §3 without risking the *basic* properties of the *Gets* reduction approach. In principle, the reduction approach proposed here can be applied to any simultaneous equation model (SEM) of the form:

$$\mathbf{B}\mathbf{y}_t = \boldsymbol{\Gamma}\mathbf{z}_t + \boldsymbol{\eta}_t, \tag{30}$$
$$\boldsymbol{\eta}_t | \mathbf{z}_t \sim \mathsf{NID}(\mathbf{0}, \boldsymbol{\Omega}),$$

where restrictions on $(\mathbf{B}, \boldsymbol{\Gamma})$ guarantee that the model is identified in sense of the famous Cowles' Commission rank condition. If the system is recursive (*i.e.* $\mathbf{B}$ is triangular), proposition 2 holds and implies the (asymptotic) efficiency of single-equation *Gets* reduction procedure. But in general single-equation reduction procedures will be inefficient.

Suppose that the VAR(p) model with (weakly) exogenous regressors:

$$\mathbf{y}_t = \boldsymbol{\nu} + \sum_{i=1}^{p} \mathbf{A}_i \mathbf{y}_{t-i} + \mathbf{D}\mathbf{x}_t + \boldsymbol{\varepsilon}_t, \quad \boldsymbol{\varepsilon}_t \sim \mathsf{NID}(\mathbf{0}, \boldsymbol{\Sigma}), \tag{31}$$

is the reduced form of the SEM:

$$\mathbf{B}\mathbf{y}_t = \boldsymbol{\delta} + \sum_{i=1}^{p} \boldsymbol{\Gamma}_i \mathbf{y}_{t-i} + \mathbf{R}\mathbf{x}_t + \boldsymbol{\eta}_t, \quad \boldsymbol{\eta}_t \sim \mathsf{NID}(\mathbf{0}, \boldsymbol{\Omega}), \tag{32}$$

where, say, for $K = 2$:

$$\mathbf{B} = \left[\begin{array}{cc} 1 & * \\ * & 1 \end{array} \right], \ \mathbf{R} = \left[\begin{array}{cc} \rho_{11} & 0 \\ 0 & \rho_{22} \end{array} \right], \ \mathbf{\Omega} = \left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array} \right].$$

The otherwise unrestricted SEM is exactly identified by the restrictions on matrix $\mathbf{R}$. Single-equation *Gets* model selection procedures can be applied using instrumental variables estimation (IVE) and IVE-based significance and encompassing tests as implemented in *PcGets*. Hereby, x_{2t} is the additional instrument for $\mathbf{y}_{2t}$ in first equation, and x_{1t} for $\mathbf{y}_{1t}$ in the second. This reduction procedure will inherent the properties of 2SLS estimation. Since the unrestricted reduced form is used for the construction of the instruments on the first stage, the procedure will loose efficiency relative to a *Gets* system procedure based on FIML estimators.

Many SEMs used in practice suffer from sparsely formulated dynamics, which can affect the structural stability and economic interpretability of the model. A useful application of the reduction approach proposed in this paper is the enrichment of the possibly misspecified SEMs as in (30) with additional, potentially omitted variables (collected to the vector $\mathbf{w}_t$), which do not interfere with the identification restrictions of the original model (*i.e.* they do not include elements of $\mathbf{y}_t$ and $\mathbf{z}_t$) and enter each equation unrestrictedly:

$$\begin{aligned} \mathbf{B}^* \mathbf{y}_t &= \mathbf{\Gamma}^* \mathbf{z}_t + \mathbf{C} \mathbf{w}_t + \boldsymbol{\eta}_t^*. \\ \boldsymbol{\eta}_t^* | \mathbf{z}_t, \mathbf{w}_t &\sim \mathsf{NID}(\mathbf{0}, \mathbf{\Omega}^*), \end{aligned} \tag{33}$$

The reduction process proposed in this paper would look for simplifications of C conditional on the unaltered structure of $\mathbf{B}^*$ and $\mathbf{\Gamma}^*$. The first step is the test of $C = \mathbf{0}$ to control the size of the procedure. This approach should help to identify omitted variables and avoid dynamic misspecifications. But the approach might also be useful for researchers who derive (30) from a theoretical model, but want to ensure the congruence of the empirical model (*i.e.* the absence of misspecifications) by introducing variables or dynamics which might be important for the analyzed data set, but are not subject of the theoretical analysis (see Flaschel and Krolzig, 2002, for an application of this approach to the modelling of US price and wage Phillips curves).

7 Conclusions

The aim of the paper was to propose and evaluate *Gets* reduction strategies for structural VAR models. We found that the recursive structural representation of linear VAR models ensures the efficiency of single-equation procedures. Next we analyzed the properties of the proposed model selection procedure, using a computer implementation based on *PcGets*, in a small Monte Carlo experiment to see if the procedure worked well, indifferently, or failed badly. The results come much closer to the first: The DGP specification was recovered from a large just-identified SVAR with anticipated size, and power close to commencing from the DGP itself. The accuracy and precision of the empirical impulse responses of the selected model surpass those of unrestricted models even if selection errors are taken into account. The feasibility of the proposed single-equation *Gets* procedure for the SVAR modelling of large macroeconomic data sets has been demonstrated. Finally, we outlined avenues for further research involving *Gets* procedures for the selection of the causal order, cointegrated VARs, and simultaneous equation models.

References

Akaike, H. (1985). Prediction and entropy. In Atkinson, A. C., and Fienberg, S. E. (eds.), *A Celebration of Statistics*, pp. 1–24. New York: Springer-Verlag.

Bernanke, B. S., and Blinder, A. S. (1992). The federal funds rate and the channels of monetary transmission. *American Economic Review*, **82**, 901–21.

Blanchard, O., and Quah, D. (1989). The dynamic effects of aggregate demand and supply disturbances. *American Economic Review*, **79**, 655–673.

Brüggemann, R., Krolzig, H.-M., and Lütkepohl, H. (2002). Comparison of model selection procedures for VAR processes. Mimeo, Humboldt–University, Berlin.

Brüggemann, R., and Lütkepohl, H. (2000). Lag selection in subset VAR models with an application to a U. S. monetary system. in Friedmann *et al.* (2001), pp. 107–128.

Christiano, L. J., Eichenbaum, M., and Evans, C. L. (1996). The effects of monetary policy shocks: Evidence from the flow of funds. *Review of Economics and Statistics*, **78**, 16–34.

Dungey, M., and Pagan, A. (2000). A structural VAR model of the Australian economy. *Economic Record*, **76**, 321–342.

Flaschel, P., and Krolzig, H.-M. (2002). Wage and price Phillips curves. An empirical analysis of destabilizing wage-price spirals. Discussion Paper, Department of Economics, University of Oxford.

Friedmann, R., Knüppel, L., and Lütkepohl, H. (eds.)(2001). *Econometric Studies - A Festschrift in Honour of Joachim Frohn*. Münster: LIT Verlag.

Hannan, E. J., and Quinn, B. G. (1979). The determination of the order of an autoregression. *Journal of the Royal Statistical Society*, **B, 41**, 190–195.

Hendry, D. F. (1995). *Dynamic Econometrics*. Oxford: Oxford University Press.

Hendry, D. F. (2000). *Econometrics: Alchemy or Science?* Oxford: Oxford University Press. New Edition.

Hendry, D. F., and Krolzig, H.-M. (2001). *Automatic Econometric Model Selection with PcGets*. London: Timberlake Consultants Press.

Hendry, D. F., and Krolzig, H.-M. (2003). The properties of automatic *Gets* modelling. Discussion Paper, Department of Economics, University of Oxford.

Hendry, D. F., and Mizon, G. E. (2000). Reformulating empirical macro-econometric modelling. *Oxford Review of Economic Policy*, **16**, 138–159.

Hoover, K. D., and Perez, S. J. (1999). Data mining reconsidered: Encompassing and the general-to-specific approach to specification search. *Econometrics Journal*, **2**, 167–91.

Johansen, S. (1995). *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*. Oxford: Oxford University Press.

Krolzig, H.-M. (2001). General-to-specific reductions in vector autoregressive processes. in Friedmann *et al.* (2001), pp. 129–157.

Lütkepohl, H. (1991). *Introduction to Multiple Time Series Analysis*. Berlin: Springer.

Omtzig, P. (2002). Automatic identification and restriction of the cointegration space. Thesis chapter, Economics Department, Copenhagen University.

Reale, M., and Tunnicliffe Wilson, G. (2002). Identification of vector AR models with recursive structural errors using conditional independence graphs. *Statistical Methods and Applications*, **11**,

forthcoming.

Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics*, **6**, 461–464.

Selva, D., and Hoover, K. D. (2002). Searching for the causal structure of a vector autoregression. Discussion paper, University of California, Davis.

Simon, H. A. (1953). Causal ordering and identifiability. In Hood, W. C., and Koopmans, T. C. (eds.), *Studies in Econometric Method*, No. 14 in Cowles Commission Monograph, C.H. 3. New York: John Wiley & Sons.

Sims, C. A. (1980). Macroeconomics and reality. *Econometrica*, **48**, 1–48.

Sims, C. A. (1992). Interpreting the macroeconomic time series facts: The effects of monetary policy. *European Economic Review*, **36**, 975–1000.

Sims, C. A., Stock, J. H., and Watson, M. W. (1990). Inference in linear time series models with some unit roots. *Econometrica*, **58**, 113–144.

Strotz, R. H., and Wold, H. O. A. (1960). Recursive versus non-recursive systems: An attempt at a synthesis. *Econometrica*, **28**, 417–421.

Swanson, N., and Granger, C. (1997). Impulse response functions based on causal approach to residual orthogonalization in vector autoregressions. *Journal of the American Statistical Association*, **92**, 357–367.

Wold, H. O. A. (1949). Statistical estimation of economic relationships. *Econometrica*, **17**, 1–21. Supplement.

Wooldridge, J. M. (1999). Asymptotic properties of some specification tests in linear models with integrated processes. In Engle, R. F., and White, H. (eds.), *Cointegration, Causality and Forecasting*, pp. 366–384. Oxford: Oxford University Press.